Decoding Disease Persistence in Pediatric Acute Lymphoblastic Leukemia One Single Cell at a Time

Kathleen Jane Imbach

A Thesis in the Field of Bioinformatics

for the Degree of Bachelor of Science in Biology

Georgia Institute of Technology

December 2020

Abstract

In order to understand the biological and molecular mechanisms underlying disease resistance to therapy in pediatric acute lymphoblastic leukemia (ALL), we performed an investigation utilizing single-cell RNA-sequencing (scRNA-seq) on the 10X Genomics Chromium platform. Bone marrow samples from seven patients were collected, four of whom exhibited measurable residual disease (MRD) after induction therapy, and three patients who did not. Cells from bone marrow tissue were extracted from each patient at the time of diagnosis, prior to treatment efforts. Leukemic cells were separated from peripheral immune cells using flow cytometry and ~1000 single cells were sequenced from each patients' cell populations. The goal of this study was to discern how the immune and leukemic cell populations and gene expression therein vary at the time of diagnosis between patients who do or do not respond to induction. Our results demonstrate a comparative increase in immune exhaustion signatures in the immune cells of MRD-positive patients, corroborating previous findings that implicate the role of exhaustion in resistant disease. We also show a discrepancy of cell cycle states in the leukemic cell compartment according to disease outcome, with an enrichment of blasts from MRD-negative patients exhibiting genetic signatures of S- and G2/M-phase.

Table of Contents

Introduction

Acute lymphoblastic leukemia (ALL) is the most common malignant disease diagnosed in children[1]. While the disease is associated with a high remission rate, the 15% of patients who relapse face diminished survival rates compared to those that do not experience recurrent ALL[2]. Currently, the most widely used treatment utilized for pediatric ALL patients is chemotherapy, a three-step process that begins with a critical phase known as induction[3]. During induction therapy patients typically receive three drugs: two chemotherapy drugs, often vincristine and asparaginase, and a corticosteroid[3]. The goal of this initial four-week stage is to achieve an absence of leukemic cells in the bone marrow or attain remission status[3]. The presence of leukemia cells, or measurable residual disease (MRD), is assessed after induction therapy, as this measure is heavily indicative of eventual patient outcome[2]. The high instance of recurrent ALL in children and its associated lower survival likelihood makes insight into the mechanisms involved in relapse of great importance to oncologic researchers[4]. Understanding the biological processes underlying patients' failure to achieve remission following induction therapy is crucial for the development of means to reduce the instance of relapsed ALL in children[4].

A recently developed technology known as single-cell RNA-sequencing (scRNA-seq) enables mass-sequencing of mRNA transcripts from many single cells at once[5]. Statistical analysis of the abundance of each transcript in individual cells allows for determination of differential gene expression, which often provides insight into the biological processes taking place in each cell[6]. This is an incredible advance over bulk

tissue RNA sequencing since it allows discrimination of both the distribution of diverse cell types and their molecular differences among patients.

Both bulk and single-cell RNA-sequencing methods have been previously employed to characterize the biological processes involved in ALL relapse. In an assessment of RNA expression in single leukemic cells in untreated controls and leukemic cells persisting in mice treated with two chemotherapy drugs, Ebinger et al (2016) found that the MRD cells had low expression of genes involved in cell cycle and DNA replication, but high expression of genes that facilitate cell adhesion[7]. However, the use of immunocompromised mouse models limits the direct application of these findings to human disease, and because only leukemic cells were analyzed, no insight into the possible ways in which other immune cells may influence disease status during therapeutic implementation could be gained[7].

Other studies have implicated the importance of expression patterns in peripheral immune cells within the bone marrow in indicating a patient's likelihood to respond to therapy. Sade-Feldman et al (2018) found that the expression of TGF7 in $CD8^+$ T cells may serve as a predictor of patients' ability to respond to immune checkpoint therapy (ICT), as might signatures of T cell exhaustion[8]. The limited focus of T cells in the work though does not permit comprehensive insight into the therapy's effect on all immune cells[8]. A scRNA-seq study of T-ALL patient bone marrow conducted by De Bie et al (2018) suggested that malignancy-inducing mutations tend to accumulate in multipotent progenitor cells, suggesting an unmet need for therapies that target these early cells[9].

Finally, Witkowski et al (2020) implicated the role of monocyte abundance in B-ALL disease progression—namely, they found that diagnosis and relapse samples

exhibited increased abundance of non-classical monocytes, and further demonstrated that high abundance of monocytes may be used to predict patient survival[10]. Furthermore, they revealed that non-classical monocytes at diagnosis and relapse showed increased expression of genes involved in vascular endothelial interactions, namely *PCAM1*, *TNFSF10*, and *CX3CR1*[10]. However, the initial separation of the bone marrow cells into $CD19^+$ B cells and $CD19^-CD45^+$ non-B cells may have compromised the ability to infer how the leukemic cell population explicitly differs from the peripheral immune cells in the bone marrow[10]. Such information is valuable, as it could potentially unveil variation in both the cell profiles and the interactions between these populations of cells in patients who do or do not respond to induction therapy.

In this pilot study, we aimed to provide additional information about the biological differences between patients who achieve remission and those who do not. Specifically, we focused on discerning discrepancies in the immune and leukemic cell profiles and gene expression therein between patients who have MRD after induction therapy and those that do not. This analysis was carried out via utilization of scRNA-seq on cells extracted and isolated from the bone marrow tissue of leukemia patients at the time of diagnosis. Flow cytometry was first done in order to separate the leukemic cells from the peripheral immune cell populations. mRNA transcripts were acquired from two samples associated with each patient: leukemia cells at diagnosis and immune cells at diagnosis. Statistical analyses were then performed on each patients' scRNA-seq data in order to characterize the cells and assess whether they exhibit significant differential gene expression, specifically comparing patients with and without MRD after the completion of induction.

Considering the immunological patterns associated with ALL disease outcome in the previous studies, we developed a three-fold hypothesis. First, we anticipated divergent cell type frequencies and gene expression profiles between patients who did vs did not achieve remission. More specifically, we expected that MRD-positive patients would have signatures of immune cell exhaustion and dormant leukemic states, while patients who achieved remission would have a higher proportion of proliferating leukemic cells. This study aims to provide further insight into how immunity and malignancy differ between therapy-responsive versus unresponsive patients at the time of diagnosis, so that such information may be applied to future work in improving clinical measures and diminishing relapse likelihood in pediatric B-ALL patients.

Methods

Extraction and Sequencing

Bone marrow samples from seven patients were collected, four of whom exhibited MRD after induction, and three patients who did not. Flow cytometry was used to separate $CD45^{+}CD10^{+}CD19^{+}$ blasts from all other viable $CD45^{+}$ cells to separate leukemic cells from peripheral immune cells. From each patient, scRNA was sequenced from ~1000 cells each from peripheral immune cells and leukemic cells before treatment. Patients' DNA was also extracted and genotyped for post-sequencing annotation of cells by individual.

Genotyping, QC and Cell Annotation

In scRNAseq analysis, transcripts are assigned three barcodes, one for the sample, one for the cell, and one UMI (unique molecular identifier) for the actual mRNA molecule. Since we pooled four individuals into one sample to save costs, a fourth barcode was required to assign cells to individuals. This barcode is actually the genotypes ascertained from the RNA sequence reads, which are compared to each individual's genotype profile with Demuxlet software[11]. Once labeled, the transcript data was processed using the Seurat package in R® statistical software. Quality control was implemented on the samples to retain cells only with transcripts for more than 200 unique genes and less than 30% mitochondrial contribution. The batched samples were then separated by individual.

All samples, both leukemic and nonleukemic, were then combined into one data object to ensure that flow cytometry had effectively separated malignant cells from peripheral immune cells. The combined data object was log-normalized and scaled using the most highly variable genes in the dataset. Principal component analysis (PCA) was performed and used in Shared Nearest Neighbor (SNN) analysis was implemented in Seurat's *FindNeighbors* function, followed by a Uniform Manifold Approximation and Projection (UMAP) dimensional reduction technique, both using a dimensions parameter of 1:35. Finally, a UMAP plot was constructed and clusters were visually separated by malignancy. The groups of non-leukemic cells that mapped with the leukemic clusters were relabeled to reflect their true malignant identities. The few leukemic cells that mapped to the peripheral immune cell regions were omitted from downstream analyses.

Cell Clustering and Characterization

Once the cells were labeled to reflect their true malignancy states, analysis of non-leukemic and leukemic cell populations were performed separately. The nonleukemic cells were analyzed first, beginning with splitting the data object according to individual. Each subject's data was normalized and scaled using Seurat's *SCTransform* function. Each individual's sample was then combined back into a combined data object using Seurat's integration workflow to account for possible batch effects in the dataset. PCA, SNN and UMAP were then run as had been done previously, again with dimensions 1:35. UMAP plots were constructed at various resolutions and the variation with the most stable cluster assignment was chosen for downstream analysis - in this case, a resolution of 9.0, which resulted in a total of 57 distinct cell clusters. After grouping these cells to these clusters, the original raw counts of the combined object were log-normalized and

scaled using the most highly variable genes. Cell-specific gene markers were assessed in each cluster through the creation of Dot Plots, which provided guidance to label the various cluster groups by cell type. After investigating the expression of cell- specific genes across the clusters, the 57 cell groups were reduced to 11 distinct cell types.

The workflow used to process the peripheral immune cells was also utilized to analyze the leukemic cell population. However, a UMAP resolution of 0.6 was found to be most stable for these cells, resulting in 10 total leukemic clusters. The original count data for these cells were also log-normalized and scaled using the most highly variable genes, however this time Seurat's *FindMarkers* function was utilized to determine the RNA transcripts positively defining each of the 10 cell groups. For each cluster, the genes with a p-adjusted (p-adj) less than 0.05 were input into functional enrichment analysis, ToppFun[12].

The *CellCycleScoring* pipeline was then run on the leukemic cells, which assigns each cell a score based on its expression of G2/M and S phase genes. The resulting assignments were then evaluated according to the composition of the cell cycle phases by disease outcome, timepoint, and leukemic cluster.

An assessment of immune cell exhaustion (loss of effector and proliferation activity) was conducted on the peripheral immune cell populations. Two methods were used to assign each cell a score based on their expression of exhaustion-specific marker genes. The first exhaustion measure was calculated by totaling the normalized expression values for each of the exhaustion genes for each cell. The second measure resulted from the summed raw number of exhaustion gene transcripts per cell, normalized by the total transcripts per cell. In each case, the distribution of scores across all cells was assessed,

and cells having upper-outlier score values in the two distributions were noted and labeled as highly exhausted cells.

Analysis of Differentially Expressed Genes

The major aim of this study was to unveil potential biological differences in the bone marrow compartment of ALL patients according to their post-therapeutic outcome. To do this, analyses were conducted evaluating the genes differentially expressed according to patient outcome across the samples. Differential gene expression was assessed using a Wilcoxon Rank Sum test between cell groups. This analysis was performed on the overall leukemic cell group as well as overall within the nonleukemic immune cell compartment to evaluate expression differences between cells from patients who did and did not have MRD post-induction. In addition to evaluating expression discrepancies according to disease outcome within these large cell groups, the same gene expression examination was conducted within each of the subpopulations comprising these cellular compartments. More specifically, outcome associated DEGs were determined within each of the leukemic cell clusters and within each of the immune cell groups. Genes were considered significantly differentially expressed if they exhibited a p-adj value less than 0.05 and a $\log_2$FC value greater than 1. Such significant genes were also input to ToppFun for functional enrichment analysis[12]. Pathways with more than 100 genes in the annotation and with a resulting Bonferroni value less than 5e-2 were specifically noted.

Results

After quality control assessments and division of single cells into their malignant and non-malignant states, a total of 18,974 single cells were retained for further analysis: 5,906 nonleukemic immune cells and 13,068 leukemic cells.

Leukemic Cell Population

After clustering, the leukemic cells were assigned to 10 distinct clusters, labeled 0-9 (Fig. 1). Each leukemic cluster was analyzed by percent contribution of cells by patients according to their disease outcome (Table 1). Clusters 5 and 8 were discovered to have a higher composition of cells from MRD-negative patients than anticipated, while cluster 9 was almost exclusively dominated by cells from MRD-positive patients. A two-sided Fischer's exact test revealed that the contingency table indicates a non-independent relationship between cluster assignment and disease outcome (p-value < 5e-4).

The genes defining each leukemic cluster were assessed, and genes with a p-adj value less than 0.05 were subjected to PANTHER's gene ontology (GO) enrichment analysis[13,14]. Cluster 5 was defined by genes associated with DNA replication and senescence evasion. Cluster 8 genes were associated with antigen processing and presentation via class I MHC and both IFN and TNF signaling pathways. Cluster 9 genes were associated with cell differentiation, inhibition of cell migration, and the p38MAPK cascade.

A cell cycle phase analysis of the leukemic cells revealed significant association with disease outcome (Pearson's Chi-squared test p-value < 2.2e-16). A higher proportion of leukemic cells from MRD-negative patients exhibited gene expression characteristic of G2/M- and S-phase when compared to leukemic cells from MRD-positive patients (Fig. 2).

Furthermore, analysis of cell-cycle phase assignment according to leukemic cluster revealed that cluster 5 is dominated by cells expressing G2M- and S-phase genes, corroborating the previous GO finding that this cluster is defined by genes characterized by DNA replication and senescence evasion pathways (Fig. 3).

The number of DEGs according to disease outcome within each of the leukemic cell clusters indicate a strong discrepancy in gene expression in clusters 0, 6, and 8 specifically (Table 2).

Functional enrichment analysis of the genes upregulated in MRD-negative patients' cells within each leukemic cluster revealed a few molecular trends in these cells. Clusters 2, 6, and 8 exhibited increased expression of genes associated with RNA binding, with genes in clusters 6 and 8 associating with the respirasome complex. Upregulated genes in cluster 0 were implicated in the AP-1 complex and B cell survival (*JUND*, *FOS*, *PIK3R1*). Cluster 5 MRD-negative associated genes were connected to negative regulation of myeloid cell differentiation (*LMO2*, *NFKBIA*, *CDK6*, *FBXW7*).

The most notable trend in pathway enrichment results from the genes upregulated in the cells from MRD-positive patients in each leukemic cluster was the recurrence of leukocyte activation, which was exhibited in all clusters but 5 and 7. Response to cytokine and immune effector process pathways also emerged often, occurring in clusters

0, 1, 3, 6 and 8. Additionally, the IgM complex and cell adhesion were correlated with clusters 0, 1, 2, 3, 4, and 5. Cluster 6 did not exhibit genes related to the IgM complex, but was associated with adhesion and the actin cytoskeleton, and both clusters 6 and 8 upregulated of various receptor genes. The MRD-positive associated genes of clusters 0, 6 and 8 also associated these clusters with genes found to be upregulated in hematopoietic progenitor cells (HPC) of B lymphocyte lineage $CD34^{+}CD45RA^{+}CD10^{+}$, as found by Haddad et al (2004)[15].

Nonleukemic Cell Population

The nonleukemic immune cells were assigned to 11 distinct cell types: cells with high mitochondrial genes (considered low-quality and omitted in downstream analyses), Naïve $CD8^{+}$ T cells, Naïve $CD4^{+}$ T cells, nucleated erythrocyte precursors, T regulatory cells, memory/effector $CD4^{+}$ T cells, B cells, memory/effector $CD8^{+}$ T cells and natural killer T cells (unable to be differentiated due to similar gene expression profiles), myeloid cells, natural killer T cells, natural killer cells, and hematopoietic stem and progenitor cells. These cells are labeled as “High_MT”, “NCD8”, “NCD4”, “Ery”, “Treg”, “MCD4”, “B”, “MCD8/NK-T”, “Myeloid”, “NK-T”, “NK” and “HSPC”, respectively, in the UMAP plot (Fig. 4).

Comparing the proportions of each cellular compartment revealed association between cell type and disease outcome (Pearson’s Chi-squared test p-value < 2.2e-16). The most pronounced contrast is the higher contribution of nucleated erythrocyte precursors in the MRD-negative cell compartment when compared to that of the MRD-positive patients.

The immune exhaustion analysis of this cell group revealed a total of 416 cells with high exhaustion scores. The majority of these cells were from MRD-positive patients (Fig. 5). Pearson's Chi-squared test with Yates' continuity correction revealed dependence between exhaustion assignment and disease outcome (p-value < 5.2e-13).

The number of significant DEGs within each of the peripheral immune cell types indicate that some immune cells may vary more according to post-induction disease outcome than others (Table 3). The cell types with the greatest number of DEGs are the various T cell groups (including the combined MCDB/NK-T compartment and Tregs) and the B cells, suggesting that lymphocyte activity may be especially important in understanding the immune differences predisposing induction response in pediatric ALL patients.

Each immune cell type exhibited almost identical gene ontology results for the genes significantly upregulated in MRD-negative patients. Each cell type (excluding erythrocytes) exhibited expression of genes associated with the molecular functions of organic acid binding, oxidoreductase activity, heme binding, tetrapyrrole binding and transporter activity. Additionally, B cells, the combined MCD8/NK-T cell group, and NCD8 cells expressed genes corresponding to lyase activity. In terms of the biological processes associated with the genes upregulated in the MRD-negative patients' immune cells, each cell type exhibited genes characteristic of cellular detoxification and antibiotic metabolic processes. Each cell type but erythrocytes, naïve T cells and NK cells also exhibited genes characterized to be associated with a response to reactive oxygen species.

While there is a lack of significant GO results for erythrocytes, myeloid cells, NCD8 cells, NK cells, NK-T cells and Treg cells within the MRD-positive cellular

compartment, patterns emerged among the pathways associated with significantly upregulated genes in therapy-unresponsive patients' other immune cell groups. B cells, HSPCs and NCD4 T cells were associated with protein translation processes and the ribosome, as well as co-translational protein targeting to the membrane and protein localization to the endoplasmic reticulum. In contrast, the significant genes upregulated in the MCD4 T cells of MRD-positive patients were linked to regulation of cell migration and motility, response to chemokines, and leukocyte homeostasis. Genes upregulated in the combined MCD8/NK-T cell group were found to be implicated in defense response, response to cytokine, leukocyte activation, and leukocyte differentiation processes.

Figures

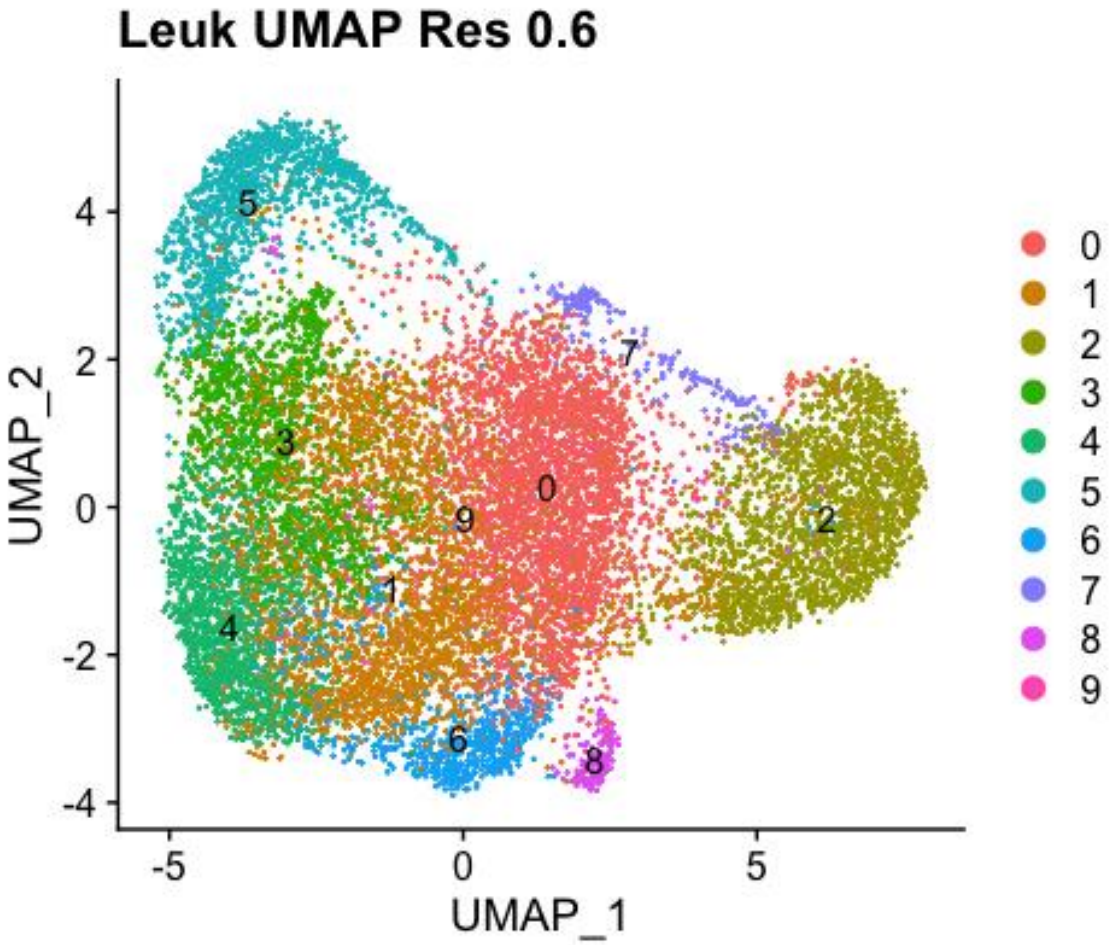

Figure 1. Dimensional Reduction Plot of Clustered Leukemic Cells.

Dimensional reduction of the leukemic cell population using UMAP with cells colored and labeled by cluster assignment.

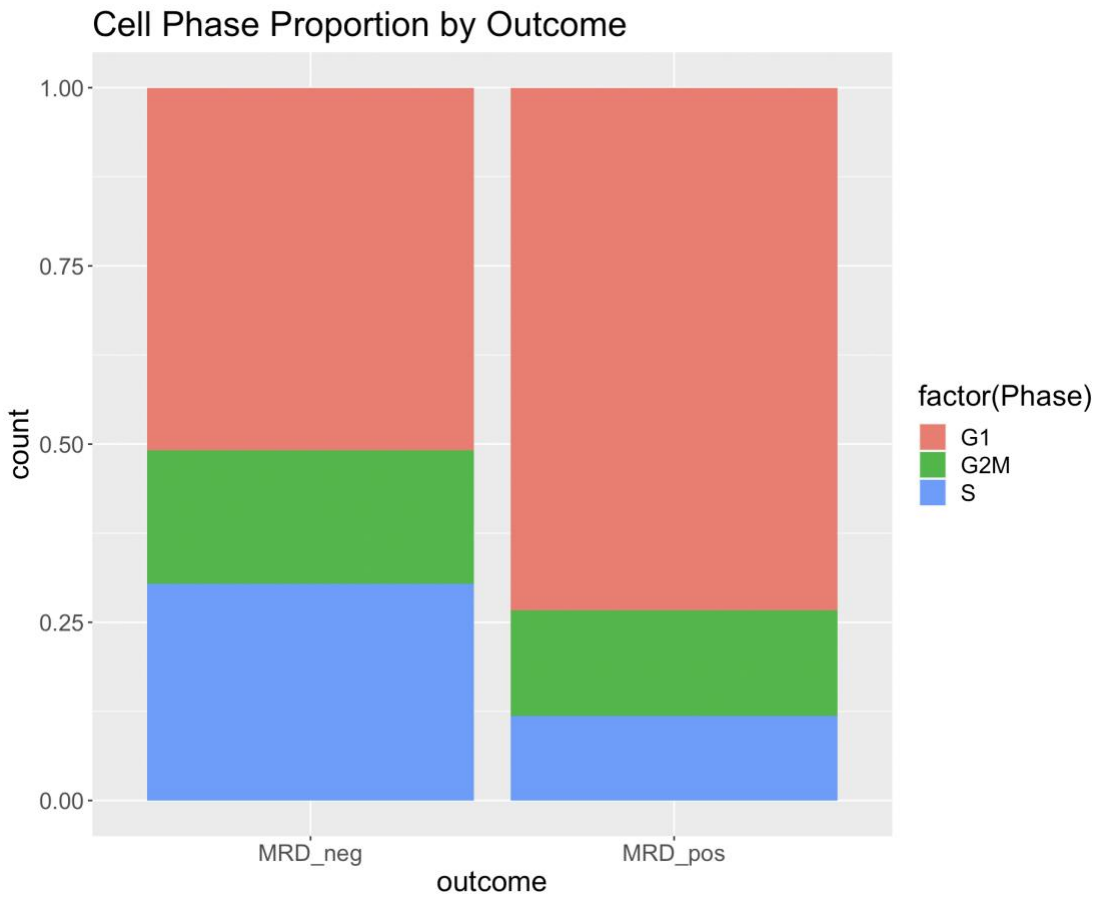

Figure 2. Leukemic Cell Cycle Phase Proportions by Outcome.

Stacked proportion plot showing proportion of cell-phase estimation of leukemic cells according to patient disease outcome.

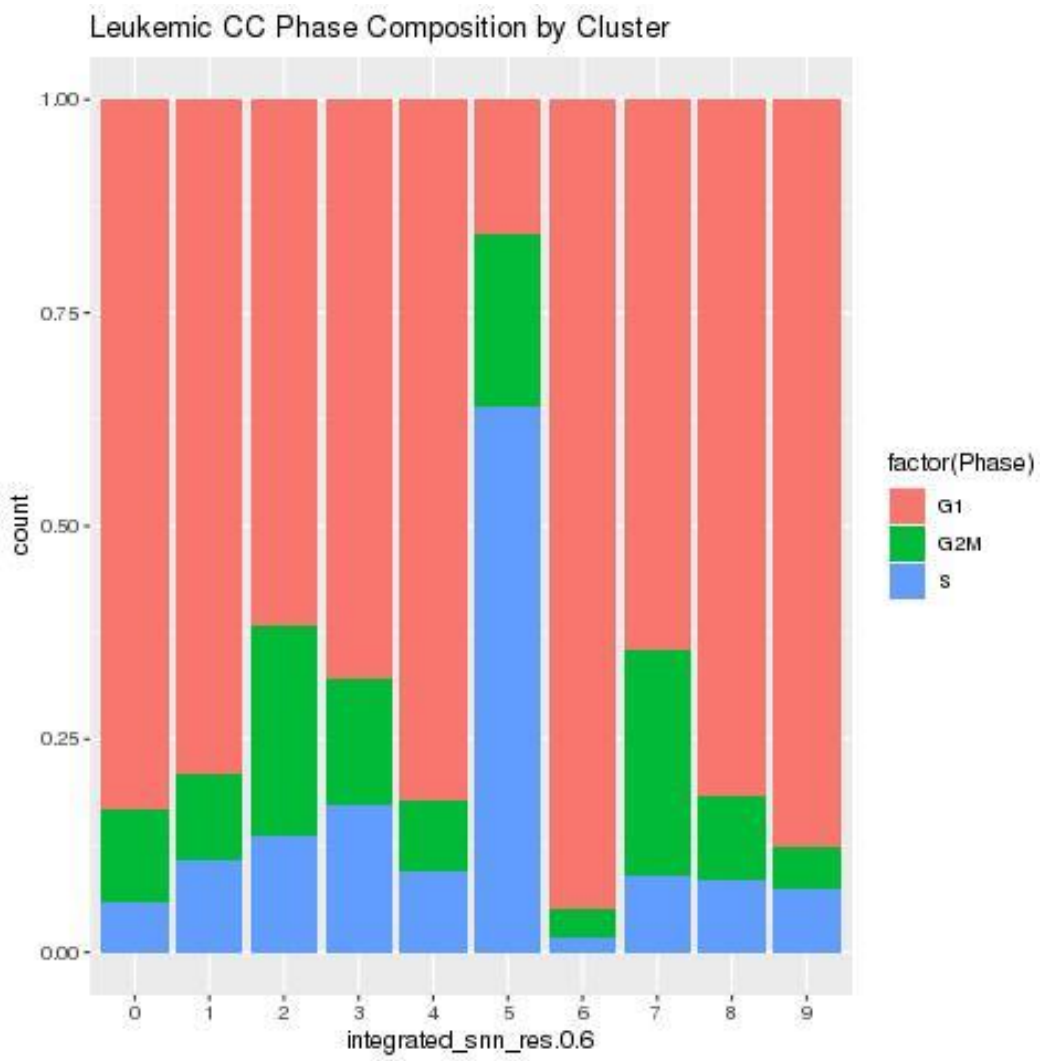

Figure 3. Leukemic Cell Cycle Phase Proportions by Cluster.

Stacked proportion plot showing proportion of cell-phase estimation of leukemic cells according to cluster assignment.

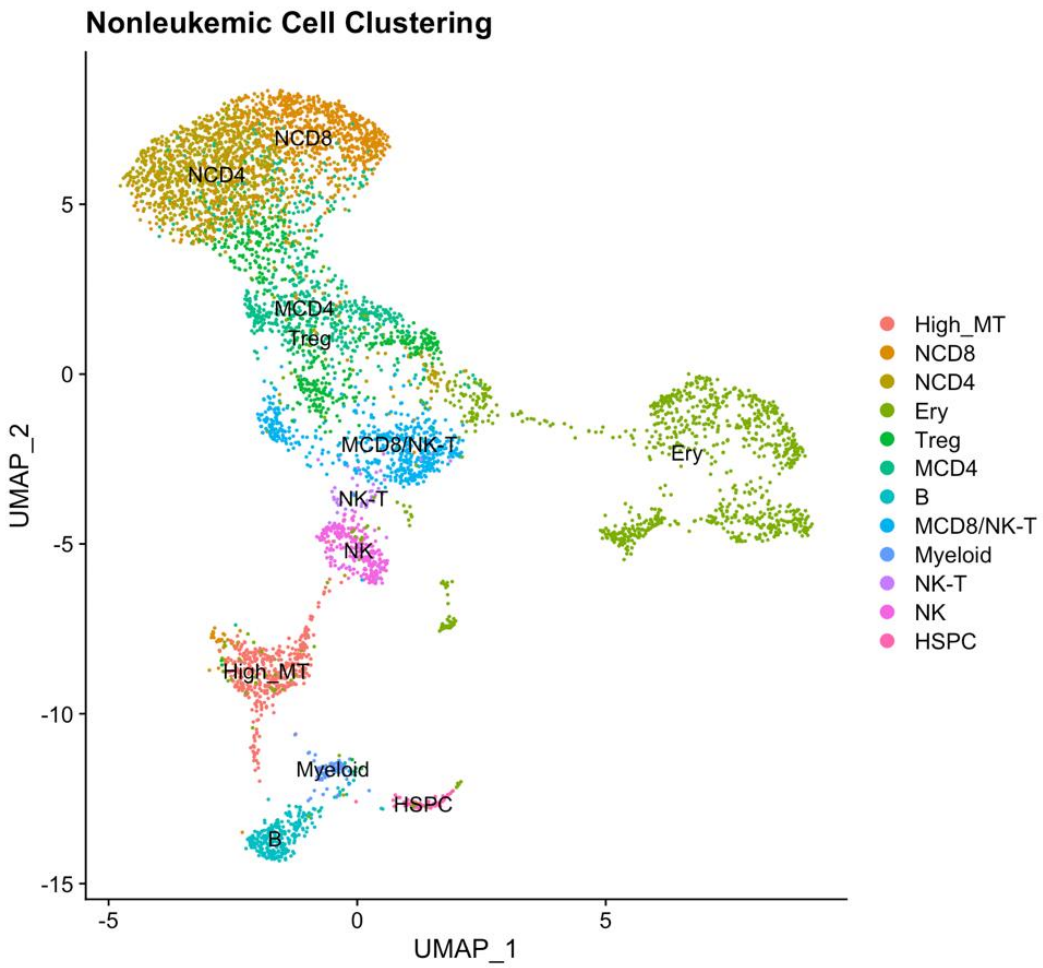

Figure 4. Dimensional Reduction Plot of Peripheral Immune Cells.

Dimensional reduction of nonleukemic cell population using UMAP with cells colored and labeled by immune cell assignment.

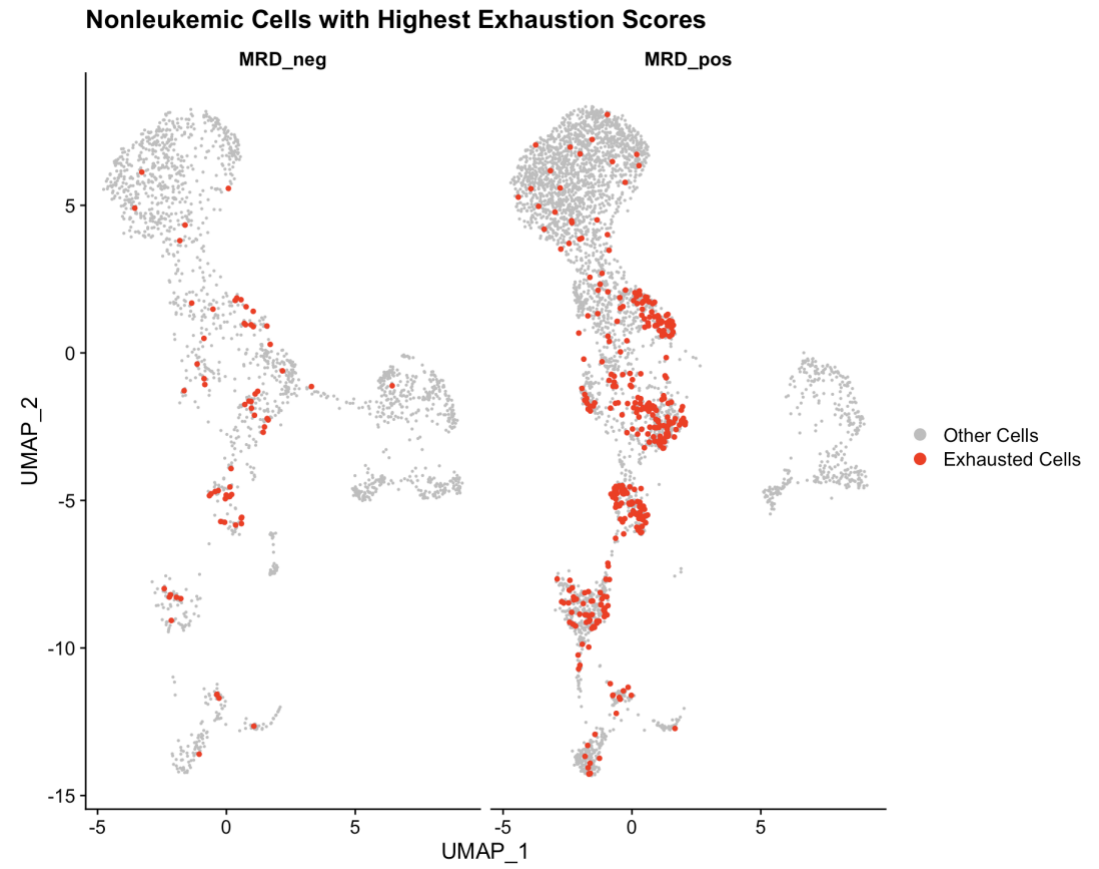

Figure 5. Nonleukemic Dimensional Reduction Plot of Exhausted Cells.

Nonleukemic UMAP plot, split by disease outcome, with high exhaustion-scoring cells highlighted.

Tables

Table 1. Leukemic Cluster Cell Contribution by Outcome.

Cluster	**# MRD negative cells**	**# MRD positive cells**	**Total**	**Proportion MRD neg to total**
0	807	2344	3151	25.6%
1	676	2153	2829	31.40%
2	236	1584	1820	14.90%
3	247	1181	1428	20.91%
4	303	1086	1389	27.90%
5	360	865	1225	**41.62%**
6	103	616	719	16.72%
7	35	176	211	19.89%
8	53	122	175	**43.44%**
9	2	119	121	**1.68%**
Total	2822	10246	13068	21.6%

Leukemic cluster cell contribution by disease outcome. Fischer's two-sided exact test resulted in p-value < 5e-4. Cluster 5 and 8 have a notably higher contribution of MRD-negative patients' cells, while cluster 9 is comprised almost exclusively of cells from MRD-positive patients.

Table 2. Leukemic Cluster DEGs by Outcome.

Cluster	MRD-Neg	MRD-Pos
0	40	78
1	32	53
2	38	24
3	31	59
4	19	34
5	27	17
6	79	86
7	29	5
8	79	69

Table displaying the number of DEGs according to post-induction disease status with $log_2FC > 1$ and p-adj < 0.05 within each leukemic cluster. Cluster 9 was not included due to its overwhelming composition of MRD-positive patients' cells.

Table 3. Immune Cell Type DEGs by Outcome.

Cell Type	**MRD-Neg**	**MRD-Pos**
B	29	27
Ery	0	4
HSPC	9	5
MCD4	14	15
MCD8/NK-T	29	76
Myeloid	7	0
NCD4	17	12
NCD8	16	11
NK	4	0
NK-T	10	0
Treg	15	14

Table displaying the number of DEGs according to post-induction disease status with $log_2FC > 1$ and p-adj<0.05 within each identified immune cell type.

Discussion

As anticipated, both the nonleukemic and leukemic cell populations exhibit divergent molecular characteristics according to disease outcome. In the case of the leukemic cell population, it was observed that MRD-negative patients' cells profiled prior to therapy are defined by active DNA replication and senescence evasion. Because induction therapy functions by targeting rapidly proliferating cells, it is intuitive that patients whose malignant cells are undergoing division will respond more effectively to this therapeutic approach. MRD-positive patients' cells, by contrast, are defined by leukocyte activation, immune effector processes, and a lower proportion of cells with genes attributed to G2M- and S-phases of the cell cycle. This observation is consistent with a proportion of cells not being actively dividing and hence able to resist chemotherapy and contribute to residual disease.

Analysis of DEGs within the leukemic cell compartment revealed yet more different trends between the malignant cells at diagnosis between patients who would then go on to either exhibit a response or resistance to induction. The RNA-binding and cellular respiration tendencies of MRD-negative leukemic cells may be due to their increased proliferation activity as suggested by the cell cycle assignment analysis. Furthermore, the implication of specific AP-1 complex proteins, most notably *JUNB* and *FOS*, may indicate the prevention of senescence or apoptosis in leukemic cells in MRD-negative patients[16], perhaps also enabling these cells to be effectively targeted by therapy. Finally, the association of cluster 5 MRD-negative upregulated genes with negative regulation of myeloid cell differentiation corroborates the findings of Witkowski et al that myeloid lineage cells may play an essential role in resistant ALL (2020)[10].

By contrast, the correlation of genes upregulated in MRD-positive leukemic cells with pathways rooted in leukocyte activation and immune effector processes may suggest that these malignant cells are chronically stimulating nearby effector cells, perhaps playing a role in the immune cell exhaustion seen in these patients' immune cells. The repetition of adhesion-related pathways in MRD-positive leukemic cells support the findings of Ebinger et al (2016) that implicate this process in resistant leukemia cells[7]. Finally, the relation of the upregulated genes in malignant cells from unresponsive patients to genes characteristic of B-lineage HPCs indicates that these cells may exhibit stem-cell-like phenotypes, allowing them to evade targeting by induction therapy.

The peripheral immune cells' molecular hallmarks also differ according to disease outcome. Firstly, the percent contribution of each immune cell type was found to vary significantly according to disease outcome, most notably with nucleated erythrocyte precursors comprising a much larger fraction of the total cells in MRD-negative patients. It was also discovered that more immune cells exhibited genetic profiles characteristic of immune exhaustion in MRD-positive patients than the cells of MRD-negative patients, further enforcing the idea that immune exhaustion may play a large role in patients' inability to respond to therapy.

Among the peripheral immune cells, those from MRD-negative patients exhibited upregulation of genes involved in detoxification and response to radical oxygen species, while the MRD-positive patients' immune cells exhibited disproportionate expression of genes implicated in protein translation processes and, in the case of MCD4 T cells and MCD8/NK-T cells, response to signaling molecules and other leukocyte-specific processes. These findings suggest that while immune cells of patients sensitive to

induction therapy elicit a strong response to the unfavorable conditions imposed by the malignant cells in the tumor microenvironment, the majority of immune cells in patients with resistant disease instead exhibit the normal cell functioning of protein production.

It was particularly interesting that various immune cells from MRD-negative patients exhibited upregulation of genes involved in heme/tetrapyrrole binding, and closer speculation of the specific genes implicated upregulation of at least one hemoglobin gene in every lymphoid cell type (*HBA1*, *HBA2*, *HBB*, *HBD*, *HBM*). Expression of such genes in non-erythroid cells is curious, but Kuo et al (2017) previously demonstrated secretion of extracellular vesicles (EV) by erythrocytes, perhaps explaining the source of these hemoglobin mRNA transcripts[17]. In the context of our findings, this could indicate that the beta-globin mRNA transcripts detected in the peripheral immune cells of MRD-negative patients may be the result of the uptake of EV secreted by red blood cells. Furthermore, work done by Danesh et al (2014) implicate that erythrocyte-derived EVs may support T-cell survival and proliferation, perhaps explaining why we observe an upregulation of HB genes in immune cells coinciding with a favorable disease outcome[18].

Overall, the results of this study provide interesting insight into the ways in which the bone marrow microenvironment may differ in patients who do versus do not respond to therapy. It may be of particular interest to future researchers to further investigate the mechanisms by which leukemic cells avoid proliferation, and thereby evade targeting by therapy, and the variables underlying whether or not a patient will experience immune cell exhaustion. It may also be useful to discern the role played by both leukocyte

activation and adhesion mechanisms in resistant leukemic cells, and how these processes might be involved in therapy evasion in resistant ALL.

Works Cited

1. Cooper, S. L. & Brown, P. A. Treatment of pediatric acute lymphoblastic leukemia. *Pediatric clinics of North America* **62**, 61-73, doi:10.1016/j.pcl.2014.09.006 (2015).

2. Ceppi, F. *et al.* Improvement of the Outcome of Relapsed or Refractory Acute Lymphoblastic Leukemia in Children Using a Risk-Based Treatment Strategy. *PloS one* **11**, e0160310-e0160310, doi:10.1371/journal.pone.0160310 (2016).

3. American Cancer Society. Treatment of Children With Acute Lymphocytic Leukemia (ALL). (2019).

4. Bhatla, T. *et al.* The biology of relapsed acute lymphoblastic leukemia: opportunities for therapeutic interventions. *Journal of pediatric hematology/oncology* **36**, 413-418, doi:10.1097/MPH.0000000000000179 (2014).

5. AlJanahi, A. A., Danielsen, M. & Dunbar, C. E. An Introduction to the Analysis of Single-Cell RNA-Sequencing Data. *Molecular Therapy - Methods & Clinical Development* **10**, 189-196, doi:10.1016/j.omtm.2018.07.003 (2018).

6. Conesa, A. *et al.* A survey of best practices for RNA-seq data analysis. *Genome Biology* **17**, 13, doi:10.1186/s13059-016-0881-8 (2016).

7. Ebinger, S. *et al.* Characterization of Rare, Dormant, and Therapy-Resistant Cells in Acute Lymphoblastic Leukemia. *Cancer cell* **30**, 849-862, doi:10.1016/j.ccell.2016.11.002 (2016).

8. Sade-Feldman, M. *et al.* Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma. *Cell* **175**, 998-1013.e1020, doi:https://doi.org/10.1016/j.cell.2018.10.038 (2018).

9. De Bie, J. *et al.* Single-cell sequencing reveals the origin and the order of mutation acquisition in T-cell acute lymphoblastic leukemia. *Leukemia* **32**, 1358-1369, doi:10.1038/s41375-018-0127-8 (2018).

10. Witkowski, M. T. *et al.* Extensive Remodeling of the Immune Microenvironment in B Cell Acute Lymphoblastic Leukemia. *Cancer Cell* **37**, 867-882.e812, doi:10.1016/j.ccell.2020.04.015 (2020).

11. Kang, H. M. *et al.* Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. (2018).

12. Chen, J., Bardes, E. E., Aronow, B. J. & Jegga, A. G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Research* **37**, W305-W311, doi:10.1093/nar/gkp427 (2009).

13. Thomas, P. D. *et al.* PANTHER: a library of protein families and subfamilies indexed by function. *Genome research* **13**, 2129-2141, doi:10.1101/gr.772403 (2003).

14. Thomas, P. D. et al. Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. Nucleic acids research 34, W645-W650, doi:10.1093/nar/gkl229 (2006).

15. Haddad, R. et al. Molecular characterization of early human T/NK and B-lymphoid progenitor cells in umbilical cord blood. Blood 104, 3918-3926, doi:10.1182/blood-2004-05-1845 (2004).

16. Weitzman, J. B., Fiette, L., Matsuo, K. & Yaniv, M. JunD Protects Cells from p53-Dependent Senescence and Apoptosis. *Molecular Cell* **6**, 1109-1119, doi:https://doi.org/10.1016/S1097-2765(00)00109-X (2000).

17. Kuo, W.P., Tigges, J.C., Toxavidis, V., & Ghiran, I. Red Blood Cells: A Source of Extracellular Vesicles. *Methods Mol Biol* 1660, 15-22, doi:10.1007/978-1-4939-7253-1_2. PMID: 28828644 (2017).

18. Danesh, A. *et al.* Exosomes from red blood cell units bind to monocytes and induce proinflammatory cytokines, boosting T-cell responses in vitro. *Blood* **123**, 687-696, doi:10.1182/blood-2013-10-530469 (2014).